

SELECTIVELY BALANCING UNIT VECTORS

AART BLOKHUIS AND HAO CHEN

ABSTRACT. A set U of unit vectors is selectively balancing if one can find two disjoint subsets U^+ and U^- , not both empty, such that the Euclidean distance between the sum of U^+ and the sum of U^- is smaller than 1. We prove that the minimum number of unit vectors that guarantee a selectively balancing set in \mathbb{R}^n is asymptotically $\frac{1}{2}n \log n$.

A set of unit vectors $U = \{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ is said to be *selectively balancing* if there is a non-trivial linear combination $\mathbf{v} = \sum \varepsilon_i \mathbf{u}_i$ with coefficients $\varepsilon_i \in \{-1, 0, 1\}$ such that the Euclidean norm $\|\mathbf{v}\| < 1$. In other words, U is selectively balancing if one can select two disjoint subsets U^+ and U^- , not both empty, such that

$$\left\| \sum_{U^+} \mathbf{u} - \sum_{U^-} \mathbf{u} \right\| < 1.$$

Note that the inequality must be strict for the problem to be nontrivial. Otherwise, one could always balance U by choosing the coefficients to be zero for all but one unit vector.

The term “balancing” refers to the classical vector balancing problems, which typically try to assign coefficients ± 1 to vectors so that the signed sum has a small norm. Various norms could be considered for balancing vectors, and different conditions can be imposed on the vectors; see e.g. [Spe77, Spe81, BG81, Spe86, Ban93, Gia97, Ban98, Swa00]. The coefficient 0 is usually not considered, despite its appearance in the powerful Partial Coloring Method (see [Bec81, Spe85] and [Mat99, §4.5, 4.6]). In this note, we try to balance unit vectors with Euclidean norms, and allow the sign to be 0. In other words, one could abandon some (not all) vectors, hence the term “selectively”.

Let $\sigma(n)$ be the minimum integer m such that any m unit vectors in \mathbb{R}^n are selectively balancing. Our main result is

Theorem.

$$\sigma(n) \sim \frac{1}{2}n \log n.$$

In other words, for any two constants $c_1 > 1/2 > c_2$, we have $c_1 n \log n > \sigma(n) > c_2 n \log n$ for sufficiently large n .

Remark. The initial motivation for our investigation is a seemingly unrelated topic: the dot product representation of cube graphs. A *dot product representation* [FSTZ98] of a graph $G = (V, E)$ is a map $\rho: V \rightarrow \mathbb{R}^n$ such that $\langle \rho(u), \rho(v) \rangle \geq 1$ if and only if $uv \in E$. It was conjectured [LC14] that the $(n+1)$ -cube has no dot product representation in \mathbb{R}^n , but was disproved by the second author [Che14]. Our construction could be modified to give dot product representations of $(cn \log n)$ -cubes in \mathbb{R}^n . See the remark at the end for the general idea.

For convenience, we take the base of the logarithm as 2. For two sets $A, B \subset \mathbb{R}^n$, $A + B$ denotes the Minkowski sum, i.e. $A + B = \{\mathbf{a} + \mathbf{b} \mid \mathbf{a} \in A, \mathbf{b} \in B\}$. We will not distinguish a set consisting of a single vector from the vector itself.

The proof of the theorem is presented in two propositions.

Proposition 1. *Let $c_1 > 1/2$ be a constant. Then for sufficiently large n , any set of $c_1 n \log n$ unit vectors in \mathbb{R}^n is selectively balancing.*

2010 *Mathematics Subject Classification.* Primary 52C07; Secondary 52A38.

Key words and phrases. Balancing vectors.

H. Chen is supported by the Deutsche Forschungsgemeinschaft within the Research Training Group ‘Methods for Discrete Structures’ (GRK 1408) and by NWO/DIAMANT grant number 613.009.031.

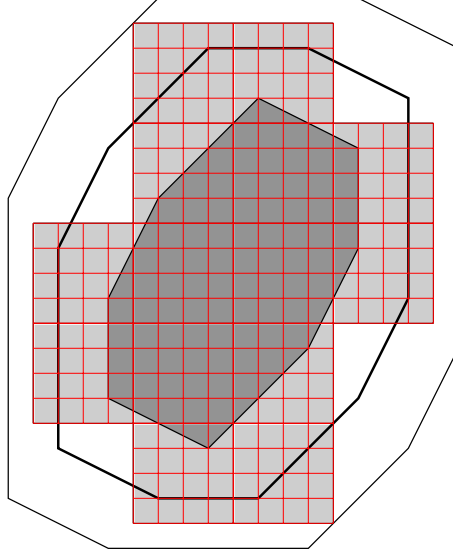


FIGURE 1. Proof of Proposition 1.

Proof. Let $Q \subset \mathbb{R}^n$ denote the unit cube $[-1/2, 1/2]^n$, and Z be the zonotope generated by $U = \{\mathbf{u}_1, \dots, \mathbf{u}_m\}$. That is,

$$Z = \left\{ \sum_{i=1}^m \lambda_i \mathbf{u}_i \mid 0 \leq \lambda_i \leq 1 \right\}.$$

In particular, Z contains all the binary combinations of U , i.e. linear combinations with coefficients 0 or 1.

Let \mathbf{v}_1 and \mathbf{v}_2 be two distinct binary linear combinations. If the Euclidean distance $\|\mathbf{v}_1 - \mathbf{v}_2\| < 1$, then $\mathbf{v} = \mathbf{v}_1 - \mathbf{v}_2$ is a non-zero linear combination of U with coefficients $-1, 0$ or 1 , and $\|\mathbf{v}\| < 1$, hence U is selectively balancing by definition. Our plan is to prove that there exist two distinct binary combinations of U at Euclidean distance < 1 if $m = c_1 n \log n$.

Let $Z^+ = Z + Q$ and $Z^{++} = Z^+ + Q$. Consider the translated unit cubes $\{Q + \mathbf{t} \mid \mathbf{t} \in \mathbb{Z}^n \cap Z^+\}$. They are contained in Z^{++} with disjoint interiors, and form a covering of Z . The number of the cubes, which is the cardinality of $\mathbb{Z}^n \cap Z^+$, is bounded from above by the volume of Z^{++} .

A zonotope can be dissected into parallelepipeds generated by linearly independent subsets of its generator; see [She74, §5] and [BR15, § 9.2]. The volume of a parallelepiped generated by unit vectors is at most 1. Since Z^{++} is (up to a translation) generated by $m + 2n$ unit vectors, its volume is at most

$$\binom{m + 2n}{n} < (e(\alpha + 2))^n,$$

where $\alpha = m/n$ and we have used Stirling's formula. We then subdivide each unit cube into $(n+1)^{n/2}$ cubes of side length $1/\sqrt{n+1}$, and estimate, very generously, at most $(e(\alpha+2)\sqrt{n+1})^n$ cubes of side length $1/\sqrt{n+1}$ with disjoint interiors. These cubes cover Z . Inside a cube of side length $1/\sqrt{n+1}$, the Euclidean distance between any two points is < 1 .

By the pigeonhole principle and the discussion before, U must be selectively balancing if $2^m > (e(\alpha+2)\sqrt{n+1})^n$. If $m = \alpha n = c_1 n \log n$ with $c_1 > 1/2$, this condition is satisfied for sufficiently large n . \square

The proof is illustrated in Figure 1.

In order to better explain our construction for the second half of the theorem, we would like to present an example first.

Example. In Figure 2 is an 5×5 integer lattice. We identify the 25 lattice points to the coordinates of \mathbb{R}^{25} . Consider two types of unit vectors: the first are the basis vectors; the second are half of

the sum of the four basis vectors corresponding to the neighbors of a lattice point. An example is given for each type in Figure 2, with the radius of the circle proportional to the component of the vector in the corresponding coordinate. There are 25 vectors of the first type, and 9 vectors of the second type, hence 34 unit vectors in total.

Consider a non-trivial linear combination \mathbf{v} of the 34 unit vectors with coefficients $-1, 0$ or 1 . If it only involves vectors of the first type, we have obviously $\|\mathbf{v}\| \geq 1$. If it only involves vectors of the second type, one verifies that \mathbf{v} has absolute value $1/2$ in at least four coordinates, hence again $\|\mathbf{v}\| \geq 1$. If both types are involved, the $1/2$'s created by vectors of the second type won't be canceled by the integers created by basis vectors, as illustrated in Figure 2. Therefore \mathbf{v} has absolute value $\geq 1/2$ in at least four coordinates, hence again $\|\mathbf{v}\| \geq 1$. We then conclude that these vectors are not selectively balancing.

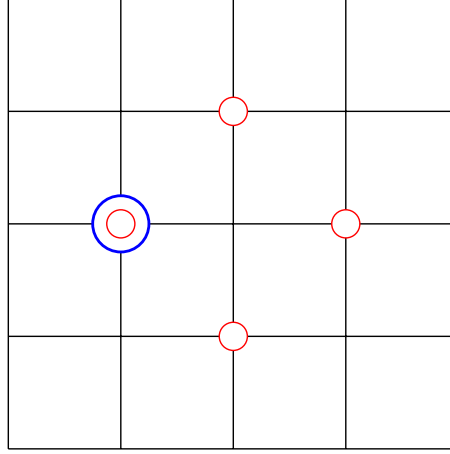


FIGURE 2. Construction of 34 unit vectors in \mathbb{R}^{25} that are not selectively balancing.

Our construction is a generalization of the example.

We need the following lemma. Let $C \subset \mathbb{R}^d$ be a set of points in strictly convex position, i.e. every point of C is a vertex of the convex hull $\text{Conv}(C)$. Let T be a finite collection of translation vectors. Then $C + T$ is the union of translated copies of C . A point $\mathbf{x} \in C + T$ is *lonely* if there is a unique $\mathbf{t} \in T$ such that $\mathbf{x} \in C + \mathbf{t}$.

Lemma. *For each $\mathbf{y} \in C$, there is a lonely point $\mathbf{x} = \mathbf{y} + \mathbf{t}$.*

Proof. Since C is strictly convex, there is a linear function $f \in (\mathbb{R}^d)^*$ such that $f(\mathbf{y}) > f(\mathbf{y}')$ for any $\mathbf{y} \neq \mathbf{y}' \in C$. We then take any $\mathbf{x} \in C + T$ that maximizes f . By construction, $\mathbf{x} \in C + \mathbf{t}$ for some $\mathbf{t} \in T$. We then have $\mathbf{x} = \mathbf{y} + \mathbf{t}$; otherwise, if $\mathbf{x} = \mathbf{y}' + \mathbf{t}$ for some $\mathbf{y}' \neq \mathbf{y}$, then $f(\mathbf{y} + \mathbf{t}) > f(\mathbf{y}' + \mathbf{t})$, contradicting the maximality of \mathbf{x} . Since the translation $\mathbf{t} = \mathbf{x} - \mathbf{y}$ is uniquely determined, \mathbf{x} is a lonely point. \square

Construction. Assume that $S \subset \mathbb{Z}^d$ contains at least 4^k integer points with the same Euclidean norm R , hence necessarily strictly convex. Let $r = \max_{\mathbf{u} \in S} \|\mathbf{u}\|_\infty$ and $L > 2r$. We now construct a set U of $(k+1)(L-2r)^d$ unit vectors in \mathbb{R}^{L^d} that is not selectively balancing.

We identify the basis vectors of \mathbb{R}^{L^d} to the integer points in $[1, L]^d$; the basis vector corresponding to $\mathbf{x} \in [1, L]^d$ is denoted by $\mathbf{e}_{\mathbf{x}}$, and $\mathbf{v}_{\mathbf{x}} = \langle \mathbf{v}, \mathbf{e}_{\mathbf{x}} \rangle$ denotes the component of \mathbf{v} in the \mathbf{x} -coordinate.

Let $S_0 \subset S_1 \subset S_2 \subset \dots \subset S_k = S$ be subsets of S such that $|S_i| = 4^i$ for $1 \leq i \leq k$. A unit vector in our set U is labeled by an integer point $\mathbf{t} \in [1+r, L-r]^d$ and an integer $i \in [0, k]$, and is defined by

$$\mathbf{u}_{\mathbf{t},i} = 2^{-i} \sum_{\mathbf{y} \in S_i} \mathbf{e}_{\mathbf{t}+\mathbf{y}}.$$

Note that, by assumption, $[1+r, L-r]^d + S \subseteq [1, L]^d$, so $U \subset \mathbb{R}^{L^d}$. We now verify that U is not selectively balancing.

Let \mathbf{v} be a non-trivial linear combination of $\mathbf{u}_{\mathbf{t},i}$ with coefficients $\varepsilon_{\mathbf{t},i} \in \{-1, 0, 1\}$.

For a fixed $i \in [0, k]$, define $\mathbf{v}_i = \sum \varepsilon_{\mathbf{t},i} \mathbf{u}_{\mathbf{t},i}$ and

$$\text{Supp}_i(\mathbf{v}) = \{\mathbf{t} \in [1+r, L-r]^d \mid \varepsilon_{\mathbf{t},i} \neq 0\}.$$

In every coordinate $\mathbf{x} \in [1, L]^d$, the component $(\mathbf{v}_i)_{\mathbf{x}}$ is an integer multiple of 2^{-i} . By Lemma , for each $\mathbf{y} \in S_i$, $S_i + \text{Supp}_i(\mathbf{v})$ has a lonely point $\mathbf{x} = \mathbf{y} + \mathbf{t}$. Therefore, in at least 4^i coordinates $\mathbf{x} \in [1, L]^d$, we have $(\mathbf{v}_i)_{\mathbf{x}} = \pm 2^{-i}$.

Define

$$j = \max\{i \in [0, k] \mid \text{Supp}_i(\mathbf{v}) \neq \emptyset\},$$

so $\mathbf{v} = \mathbf{v}_0 + \dots + \mathbf{v}_j$. For each $i < j$, $(\mathbf{v}_i)_{\mathbf{x}}$ is a multiple of 2^{-i} , hence a multiple of 2×2^{-j} , in every coordinate $\mathbf{x} \in [1, L]^d$. Since $(\mathbf{v}_j)_{\mathbf{x}} = \pm 2^{-j}$ in at least 4^j coordinates, we have $|\mathbf{v}_{\mathbf{x}}| \geq 2^{-j}$ in these coordinates. So $\|\mathbf{v}\| \geq 1$.

The second half of the theorem is proved by adjusting the parameters in the construction.

Proposition 2. *Let $c_2 < 1/2$ be a constant. Then for sufficiently large n , there are $c_2 n \log n$ unit vectors in \mathbb{R}^n that are not selectively balancing.*

Proof. Let $D = 2^d$. Consider the $(2D+1)^d$ integer points in $[-D, D]^d$. Their squared Euclidean norms are at most dD^2 . By the pigeonhole principle, there is a set $S \subset [-D, D]^d$ consisting of lattice points with the same Euclidean norm $R \leq D\sqrt{d}$, whose cardinality

$$|S| \geq \frac{(2D+1)^d - 1}{dD^2} > 4^{(d^2-d-\log d)/2}.$$

Note that $r = \max_{\mathbf{u} \in S} \|\mathbf{u}\|_{\infty} \leq D$. We have constructed m unit vectors in \mathbb{R}^n that are not selectively balancing, where $n = L^d$ and

$$\begin{aligned} m &> \frac{1}{2}(d^2 - d - \log d)(L - 2r)^d \\ &\geq \frac{1}{2}(d^2 - d - \log d)(L - 2D)^d \sim \frac{1}{2}d^2(L - 2D)^d. \end{aligned}$$

For a constant $\lambda > 1$, we take $L = \lfloor D^\lambda \rfloor = \lfloor 2^{\lambda d} \rfloor$, which is eventually bigger than $2D$. Then $n \sim 2^{\lambda d^2}$ and $\log n \sim \lambda d^2$, hence

$$\lim_{d \rightarrow \infty} \frac{m}{n \log n} > \frac{1}{2\lambda}.$$

We conclude that, as long as $c_2 < 1/2\lambda$, there are more than $c_2 n \log n$ unit vectors that are not selectively balancing for sufficiently large integers of the form $n = \lfloor 2^{\lambda d} \rfloor^d$.

If n is sufficiently large, we can always find an integer d and a constant $\mu \in (\sqrt{\lambda}, \lambda)$ such that $\lfloor 2^{\mu d} \rfloor^d \leq n \leq (\lfloor 2^{\mu d} \rfloor + 1)^d$. Hence for any $c_2 < 1/2\lambda$, we have $\sigma(n) > c_2 n \log n$ for sufficiently large n . This finishes the proof since we can choose λ to be arbitrarily close to 1. \square

Remark. We give credit to the anonymous referee for this choice of S , which helped improving the proposition. In a preliminary version, we used another S , and proved for any $c_2 < 1/3e^2$ that $\sigma(n) > c_2 n \log n$ for infinitely many n .

Remark. In the construction, we could also replace 2 by any integer $p > 2$ and, correspondingly, 4 by p^2 . In particular, if we take an *odd* integer $p \geq 5$, we obtain a set of unit vectors that is *strictly* not selectively balancing: a linear combination \mathbf{v} of U with coefficients $-1, 0$ or 1 has Euclidean norm 1 only if $\mathbf{v} \in \pm U$. Our current proof for this fact is however too long and does not fit into this short note.

Note that m unit vectors in \mathbb{R}^n that are *strictly* not selectively balancing imply a ball packing in \mathbb{R}^n whose tangency graph is a m -cube. Then we conclude the following by Proposition 5 of [KLMS11]: there is a constant c such that, for infinitely many n , the $(cn \log n)$ -cube admits a dot product representation in \mathbb{R}^n . This is actually the initial motivation of our investigation.

REFERENCES

- [Ban93] W. Banaszczyk, *Balancing vectors and convex bodies*, Studia Math. **106** (1993), no. 1, 93–100.
- [Ban98] W. Banaszczyk, *Balancing vectors and Gaussian measures of n -dimensional convex bodies*, Random Structures Algorithms **12** (1998), no. 4, 351–360.
- [Bec81] J. Beck, *Balancing families of integer sequences*, Combinatorica **1** (1981), no. 3, 209–216.
- [BG81] I. Bárány and V. S. Grinberg, *On some combinatorial questions in finite-dimensional spaces*, Linear Algebra Appl. **41** (1981), 1–9.
- [BR15] M. Beck and S. Robins, *Computing the continuous discretely*, Second, Undergraduate Texts in Mathematics, Springer, New York, 2015. Integer-point enumeration in polyhedra, With illustrations by David Austin.
- [Che14] H. Chen, *Ball packings and Lorentzian discrete geometry*, Ph.D. Thesis, 2014.
- [FSTZ98] C. M. Fiduccia, E. R. Scheinerman, A. Trenk, and J. S. Zito, *Dot product representations of graphs*, Discrete Math. **181** (1998), no. 1-3, 113–138.
- [Gia97] A. A. Giannopoulos, *On some vector balancing problems*, Studia Math. **122** (1997), no. 3, 225–234.
- [KLMS11] R. J. Kang, L. Lovász, T. Müller, and E. R. Scheinerman, *Dot product representations of planar graphs*, 2011, pp. Paper 216, 14.
- [LC14] B.-J. Li and G. J. Chang, *Dot product dimensions of graphs*, Discrete Appl. Math. **166** (2014), 159–163.
- [Mat99] J. Matoušek, *Geometric discrepancy*, Algorithms and Combinatorics, vol. 18, Springer-Verlag, Berlin, 1999. An illustrated guide.
- [She74] G. C. Shephard, *Combinatorial properties of associated zonotopes*, Canad. J. Math. **26** (1974), 302–321.
- [Spe77] J. Spencer, *Balancing games*, J. Combin. Theory Ser. B **23** (1977), no. 1, 68–74.
- [Spe81] J. Spencer, *Balancing unit vectors*, J. Combin. Theory Ser. A **30** (1981), no. 3, 349–350.
- [Spe85] J. Spencer, *Six standard deviations suffice*, Trans. Amer. Math. Soc. **289** (1985), no. 2, 679–706.
- [Spe86] J. Spencer, *Balancing vectors in the max norm*, Combinatorica **6** (1986), no. 1, 55–65.
- [Swa00] K. J. Swanepoel, *Balancing unit vectors*, J. Combin. Theory Ser. A **89** (2000), no. 1, 105–112.

E-mail address: a.blokhuis@tue.nl

E-mail address: hao.chen.math@gmail.com

DEPARTEMENT OF MATHEMATICS AND COMPUTER SCIENCE, TECHNISCHE UNIVERSITEIT EINDHOVEN